

STRENGTHENING THE VALIDITY OF SOFTWARE PROCESS IMPROVEMENT MEASUREMENTS THROUGH STATISTICAL ANALYSIS: A CASE STUDY AT ERICSSON AB

Daniél José Rocha, IT University of Göteborg, rocha.daniel.j@gmail.com

Abstract—Measuring Software Process Improvements is a challenge for software organizations. It is difficult to find a non-questionable way to show that improvement actually has taken place. Through process employee surveys it is possible to design questions and key ratios that help us understand current situations and trends. The telecom company Ericsson AB has been working with such measurements since early 2001. In this thesis we investigate the result of such a survey through basic statistical analysis. We conclude that the statistical analysis of the process employee survey investigated in this thesis shows strong consistency and validity of both survey questions and key ratios. We also conclude that we can strengthen the statistical validity in SPI measurements through the use of basic statistical analysis. In this study we suggest a set of practical steps that can be performed on SPI measurements in order to strengthen their validity. We also present an iterative process model that could be used at Ericsson to assure quality in their SPI measurements and improve future SPI measurements.

Index Terms — Software Process Improvement, Statistical Analysis, Measurement

INTRODUCTION

Since the early 1990's an increasing number of organizations has embraced the idea of Software Process Improvement (SPI). Today SPI work is generally accepted in the software engineering field as a necessity to stay competitive on the market (Dickson, G. *et al.* 1980; Hartog *et al.* 1986; Niederman *et al.* 1990; Brancheau *et al.* 1994-1995). The potential benefits for organizations successful in their SPI initiatives are huge. As reported by Herbsleb *et al.* (1994) the return on investment in SPI could potentially be as high as 500-700%. However, the failure rates in such initiatives are just as high, about 80% according to Zackrisson and Freedman (2003). The Software Engineering Institute reports equally dim results with a failure rate of about 70%. It is therefore interesting and important to ask why so many SPI initiatives fail and consequently, the importance of understanding what can be done to minimize the risk of failure in future undertakings.

In the software engineering field the area of software process improvement is well treaded. The performing of adequate measuring of such initiatives are however not as common. In order for SPI initiatives to be useful its effects need to be measured and analyzed. Dybå (2000) performed a large literature review including 55 software companies; their research shows that concern for measurement is one out of six key factors for successful process improvement activities. Measuring SPI work is

required in order to make sense of, act upon, and follow up progress. The importance of recording and measuring change is as crucial as it is complicated. Iversen and Ngwenyama (2006) discuss the complex task of measuring the effectiveness of SPI initiatives.

In order to fully understand the effects of SPI initiatives in an organization, the need to measure such initiatives has been shown to be imperative. This can be done by gathering feedback and acting according to the result of it. The feedback data can take many different forms ranging from; interviews with employees affected by change, to calculating how many less lines of code a programmer writes with the introduction of a new set of coding conventions, to process employee surveys sent out in an organization to assess what the views on current processes are among the employees. The latter is just what the telecom company Ericsson AB has done as part of making SPI happen (Börjesson and Mathiassen, 2003).

Change agents at Ericsson's R&D unit have been actively working with processes in their SPI initiatives during a seven year period. As part of these initiatives they regularly conduct a process employee survey, currently containing twelve questions related to processes knowledge and process use within the organization. The results from this survey are then compiled internally by Ericsson change agents. Improvement areas and activities are identified, compiled, and sent out to units within the organization, to be acted upon in the SPI iteration that follows. This has allowed change agents within Ericsson to get a result that can be interpreted in terms of correlation between answers to survey questions and knowledge and use of processes within the organization. The suggested improvements and the potential gain of these improvements are, however, directly dependant on how the survey results are analyzed. Bryman and Cramer (1994) states that, "an awareness of quantitative data analysis greatly enhances the ability to recognize faulty conclusions or potential biased manipulation of the information". It is important not only to measure the progress of improvement initiatives but also to perform trustworthy analysis of measurement data. This allows for an organization to present results that are accepted by the people affected by the change.

For the statistical analysis conducted in this paper of the process employee survey data we used several different statistical methods in SPSS¹, a statistical and data management package, as well as SmartPLS², a structural model based tool. With these tools we wanted to investigate the possibility of performing statistical analysis of SPI measurements to assess the potential of creating validity in them. The intention with these methods was also to address the general purpose of this research:

Assuring quality of results from SPI measurements through statistical analysis.

The rest of the paper is organized as follows. The next chapter describes the theoretical context for the study, followed by a description of the research methodology. Succeeding that, the results from the research is presented. Finally, we discuss and conclude the result.

THEORETICAL CONTEXT

Knowledge about SPI is one key factor to successful SPI initiatives and there exists a vast body of knowledge about Change Management and SPI in the field today³. Many of the known researchers in SPI literature agrees on the necessity of measuring to understand and improve practice (Humphrey, 1989; McFeeley, 1996; Grady, 1992; 1997; Weinberg, 1993; Zahran, 1997). It has been shown that measuring SPI progress is an important key factor to SPI success. However, the measurements of change are not in them self sufficient enablers of improvement and acceptance of change. Zmud (1984) states that change is most likely to happen when, a need and the mean to resolve that need, co-

¹ SPSS 16.0.1 for Windows

² SmartPLS 2.0, release M3 for Windows

³ A simple search on <http://scholar.google.com> returns around 76 000 articles on "Change Management" and more than 10 000 articles on "Software Process Improvement"

exist – this is identified as the so-called “push-pull” theory. According to Börjesson and Mathiassen (2004) process push is performed by the SPI practitioners, and practice pull by the software engineers. When these inducements exist at the same time, a change is more likely to occur.

However, if no practice pull is present, the practitioners have to be convinced to believe in the benefits that a change could have for them (Weinburg, 1997). One of the most respected theories in the domain of technology acceptance, Davis’s Technology Acceptance Model (1989) explains this: “When a person regards a process or a tool as being useful in his or her work, as well as easy to use, chances are good that the new item will be used”. One way of achieving this, is by strengthening the statistical validity of the SPI measurements within the organization by assessing the assessments.

The following subsections present SPI Measurements and Basic Statistical Analysis.

SPI MEASUREMENTS

Measuring SPI is expensive and requires a great deal from people that usually already has a lot to do, argues both Humphrey (1989) and Weinberg (1993). The complex task of measuring is widely known in the software engineering field, and in the literature many warning flags are raised when it comes to measuring SPI.

In the capability maturity model (CMM) (Paulk *et al.* 1995) we see that even though measurement is an integral part of the model, it is not until level four that the key practices area, measurement, is introduced. To reach this level an organization requires having measurements for productivity and quality for the most important software project activities across all projects as a part of an organizational measurement program (Paulk *et al.* 1995). However, not many organizations manage to actually reach CMM level four (SEMA, 2002). One explanation for the low success rate in the CMM model might be that many measurement activities in SPI focus on the end results, like increase or decrease in productivity, whereas SPI is a continuous iterative process. This process cannot exist without continuous and systematic monitoring and measurement of an organizations own processes (Zahran, 1997).

Despite many of the known difficulties when it comes to measuring SPI, we can not find any literature that argues for a halt in measurements. On the contrary, we must not stop measuring states Humphrey (1989). However, DeVellis (2003) states that if a poor measure is the only one available, the costs of using that measure may turn out to be greater than any benefits attained. This shows the importance of not only measuring SPI progress, but also performing appropriate analysis of the measurements. Statistical analysis of SPI measurements could potentially improve their validity and quality. For without proper measurements and analysis, suggested actions based on SPI outcome are at the risk of being perceived as yet another opinion (Börjesson, 2006). Yet there exists no comprehensive measurement framework in the SPI literature (Iversen and Ngwenyama, 2006), which complicates matters when it comes to actually measuring SPI progress.

BASIC STATISTICAL ANALYSIS

There are a number of powerful tools available that can assist a researcher in performing statistical analysis of data. In our study we primarily used SPSS and SmartPLS. SPSS is a tool that provides a wide variety of statistical methods for analyzing data. We used SPSS in the initial analysis and SmartPLS when building, running, and validating the process model. The theories behind these statistical methods are presented below.

DISTRIBUTION ANALYSIS

Determining the distribution of a response data set is an important and common first step in many statistical analysis undertakings. Distribution analysis is performed in order to understand the spread of the population investigated, but also important since several other statistical methods are dependent on the type of distribution represented by the data. Normal distribution is the most common and well

known distribution and several other statistical methods rely on an assumption of normality.

Normal is used to describe a symmetrical, bell-shaped curve, which has the greatest frequency of scores in the middle, with smaller frequencies towards the extreme (Gravetter and Wallnau, 2004).

By using graphical methods such as histograms with normal distribution curves and normal quantile-quantile plots (Q-Q plot), one can get a good indication if the distribution can be assumed to be normal. In the normal Q-Q plot diagram the observed values of a single variable are plotted against the expected values if the sample were from a normal distribution. Should the points cluster around a reasonably straight line, the sample can be assumed to be from a normal distribution (see Fig. I).

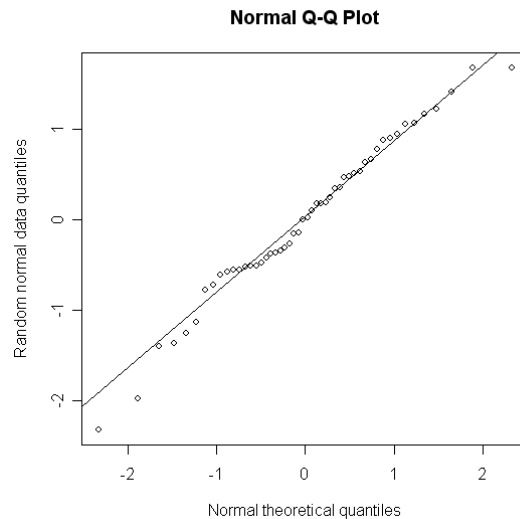


Fig. I: An example of a normal distribution presented with the graphical method, Normal Q-Q plot

There is also a set of different numerical methods for investigating what type of distribution corresponds best to the data. One numerical method for testing the assumption of normality is the One-Sample Kolmogorov-Smirnov test (K-S test). This test assesses the normality of the distribution of scores where a non-significant result (Sig. value of more than 0,05) indicates normality. The K-S test is a powerful test for assessing normality; however, it is not fitting for large populations.

Should the distribution of the data prove not to be well enough normal the statistical methods that require normal distribution are reasonably tolerant of violation of this assumption. Moreover, with a large enough population size, most literature states that, one can assume normality in the data sample when the population size >30 . However, in recent research numbers such as >250 has been suggested as more appropriate (Marasinghe *et al.* 1994). So in effect, should a large enough population exist then an assumption of normality should not cause any major problem, this is the essence of the Central Limit Theorem (CTL) (Pólya, 1920; Bernstein, 1945).

CORRELATION ANALYSIS

For correlation analysis there are a number of different statistical analysis methods that can be used. How well each method fits the data set depends on the nature of the data. One such method is the Spearman's rank order correlation coefficient or more commonly referred to as *Spearman's rho*, and reported as R_s . This is a nonparametric statistical method that can be used with ordinal data and that assumes normally distributed scores. It measures the strength of the linear relationship between two variables. It can assume values $-1 \leq R_s \leq 1$, $R_s = +/-1$, only if the relationship is perfectly linear does it assume the values 1 or -1. The sign in front indicates whether there is a positive correlation (as one variable increases, so too does the other) or a negative correlation (as one variable increases, the other decreases). Cohen (1988) recommends these threshold values for Spearman's rho coefficient (see Table I).

Table I: Recommended threshold values for Spearman's rho

Weak	$R_s = 0,10$ to $0,29$ or $R_s = -0,10$ to $-0,29$
Medium	$R_s = 0,30$ to $0,49$ or $R_s = -0,30$ to $-0,49$
Strong	$R_s = 0,50$ to $1,00$ or $R_s = -0,50$ to $-1,00$

Correlations between items can be measured to investigate if the items are related to one and other and if so how strongly. A score of zero means that there is no linear correlation between the variables and the variables might then be unrelated. This would show that some variables chosen for analysis might not fulfill the intention of the measurement.

FACTOR ANALYSIS

Factor analysis takes a large set of variables and tries to find a way to explain these variables using a smaller set of factors or components. Factor analysis is not a definitive method and it is at large up to the researcher to decide if the proposed construct for each measurement should fit on that component or in fact on a different one. It does however provide some insight into what underlying factors could explain the measurement variables.

Before performing a factor analysis there are two main issues to consider in determining whether the data is suitable for factor analysis; population size (i.e. the number of respondents in the case of a survey), and the strength of the relationship between the measurement variables (i.e. Spearman's rho). The general perception among experts is: the larger the population the better. Tabachnick and Fidell (2007) suggest that "it is comforting to have at least 300 cases for factor analysis". On the second issue, the strength of the relationship, Tabachnick and Fidell (2007) recommends correlation scores greater than 0,30. There are also two statistical methods that are of interest when determining if factor analysis is to be considered; Bartlett's test for sphericity (Bartlett, 1954), and the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy (Kaiser, 1970; 1974). Bartlett's test of sphericity should be significant ($p < 0,05$). The KMO index ranges from 0 to 1, with 0,60 as suggested threshold value for a good factor analysis, Tabachnick and Fidell (2007).

The relationships that are extracted in a factor analysis can be investigated further by building a model and testing it with the statistical method Partial Least Squares-regression (PLS).

PARTIAL LEAST SQUARES (PLS) REGRESSION ANALYSIS

PLS regression is a recent technique with the goal to predict or analyze a set of dependent variables or latent constructs from a set of independent variables or indicators (Chin, 1996). PLS regression suits survey data well as it is able to model underlying constructs even if the data should prove not to be normally distributed. There are a number of tools available to build a model with the components extracted in a factor analysis. These tools use the PLS regression for analysis. The technique models both the measurement model and the structural model, also referred to as outer versus inner model.

The measurement model involves the relationship between latent constructs (i.e. components extracted in the factor analysis) and their indicators (i.e. the variables analyzed in the factor analysis), while the structural model concerns the relationships between the latent constructs, see Fig. II.

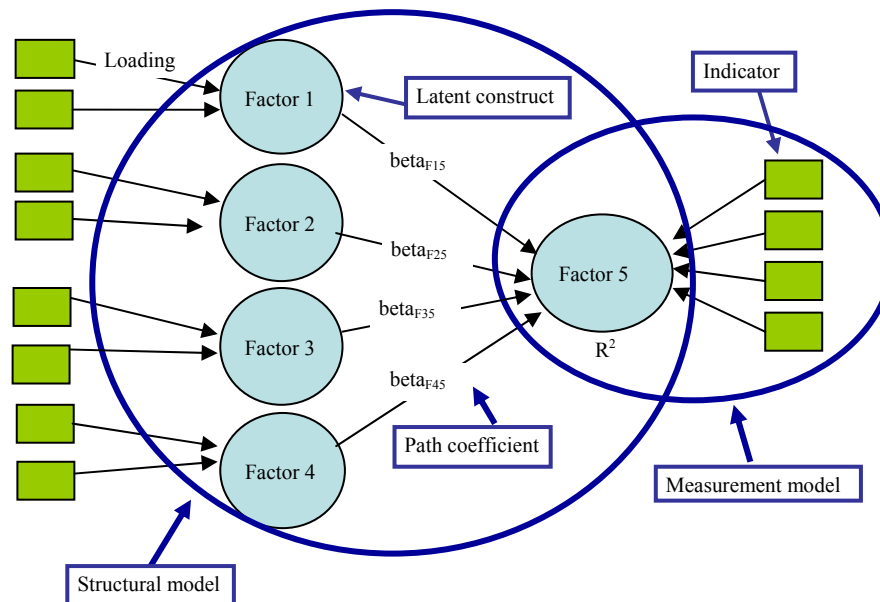


Fig. II: Explanation of concepts in PLS regression analysis

The measurement model together with the structural model will be referred to as the process model. The rest of this chapter presents the statistical methods and corresponding coefficients that are available in PLS regression analysis for explaining the consistency and validity of the Measurement Model and the Structural Model respectively.

THE MEASUREMENT MODEL

By testing the measurement model it can be determined how well the indicators explain their respective construct. What is tested is the reliability and internal consistency of the model. To assess the measurement model, a number of coefficients that can be calculated, such as:

Loading: Loading is the correlation coefficient between latent constructs and their indicators. It can be used to compute Average Variance Extracted, Composite Reliability, and Communality. The suggested threshold value for the Loading coefficient is 0,80 (Chin, 1995).

Average Variance Extracted: AVE is a statistic involving the percentage error variance in a measure. It is used for assessing the internal consistency in the measurement model. Suggested threshold values for strong AVE value is a score above 0,50 (Dillon and Goldstein, 1984).

Composite reliability: Composite reliability gives an indication to how well each of the constructs in the measurement model is described by their indicators. Recommended threshold is 0,70 (Chin, 1998). Over threshold results imply that each construct is well described by its indicators.

Cronbach Alpha: Cronbach's Alpha is a coefficient of reliability and consistency and it measures the internal consistency of the model. In other words it measures how well a set of indicators explain a single latent construct. Ideally the Cronbach Alpha score should be above 0,70 (DeVellis, 2003).

Communality: Communality is the squared correlation between one indicator and its corresponding latent construct. It measures the capacity of the indicator to describe the related latent construct. The recommended threshold for communality is 0,50 (Chin, 1998).

Convergent Validity: Convergent validity deals with examination of the degree to which the indicators of a latent construct converge with the constructs they are supposed to come together on. Factor loadings that are greater than 0,70 are considered to be a strong assessment of convergent

validity. In addition, the indicators should load stronger on their own latent construct than on the others (Trochim, 2006).

Discriminant Validity: Discriminant validity examines the degree to which the constructs diverge from each other. When the square root of the AVE for each construct is greater than their correlation with the other constructs, this indicates that they do measure different concepts (Chin, 1998). Demonstrating convergent and discriminant validity has been called measure validation (Heeler and Ray, 1972).

THE STRUCTURAL MODEL

The structural model concerns the relationships between the latent constructs. Testing the structural model determines whether there is empirical evidence for the hypothesized relationships between the constructs. There are a number of coefficients that can be calculated, such as:

Path Coefficient: Path coefficient, *beta*, which is equal to spearman's rho correlation coefficient (see Correlation Analysis), but is used to describe the linear relationship between constructs in a structural model. The significance level, *p*, of the path coefficient can be retrieved by looking up the t-statistic in a t-table⁴.

Significance Level: After an estimation of a coefficient, the t-statistic for that coefficient is the ratio of the coefficient to its standard error. That ratio can be tested against a t-table and the score determines which significance level the coefficient has, thereby indicating how much one can rely on the result obtained.

Coefficient of Determination: To calculate how much of the variance two variables share, the coefficient of determination, R^2 , can be used. This is calculated by squaring the path coefficient. For example, two variables that correlate $beta=0,70$ share $0,70*0,70=0,49$, which is to say that the independent variable is said to explain 49 percent of the dependant variable.

RESEARCH METHODOLOGY

This chapter is divided into three sections explaining how the study was conducted. Design and Procedure explains how the research and empirical material was collected and analyzed. The second section, Participants, presents the participants involved in the study. Data Sources concludes the chapter and presents which sources of data have been used in the study.

DESIGN AND PROCEDURE

The research was carried out as a case study to practically investigate a particular phenomenon within a specific empirical context and support the investigation with theory. Yin (1984) defines the case study research method as “an empirical inquiry that investigates a contemporary phenomenon within its real-life context; when the boundaries between phenomenon and context are not clearly evident; and in which multiple sources of evidence are used”.

The basis for the study was the survey results data from the employee process survey conducted in November 2007 at Ericsson. At the outset, a literature review was conducted in order to narrow in on the problem domain and formulate a research question that would contribute to the body of knowledge of SPI. Additional literature studies were conducted to acquire the necessary statistical knowledge. The statistical methods learned were then applied to the survey data, which generated a number of findings. These findings and the general research question was thoroughly analyzed in continuous participatory iterative discussions with SPI change agents and employee process survey experts at

⁴ One version of a t-table can be found at [HyperStat Online Contents](#)

Ericsson. The ideas from these discussions were then applied in further statistical analysis, as part of an iterative process. When reliable statistical findings had been extracted, a statistical expert, without affiliation to Ericsson, was contacted to discuss and validate the findings. In the end, the iterative discussions and literature material served as the basis for the conclusions drawn as to the possibility of assuring quality in SPI measurements within an SPI organization through statistical analysis. As an additional data source the author had access to Ericsson internal documentation and previous survey data and results. A conceptual image of the research methodology design can be seen in Fig. III.

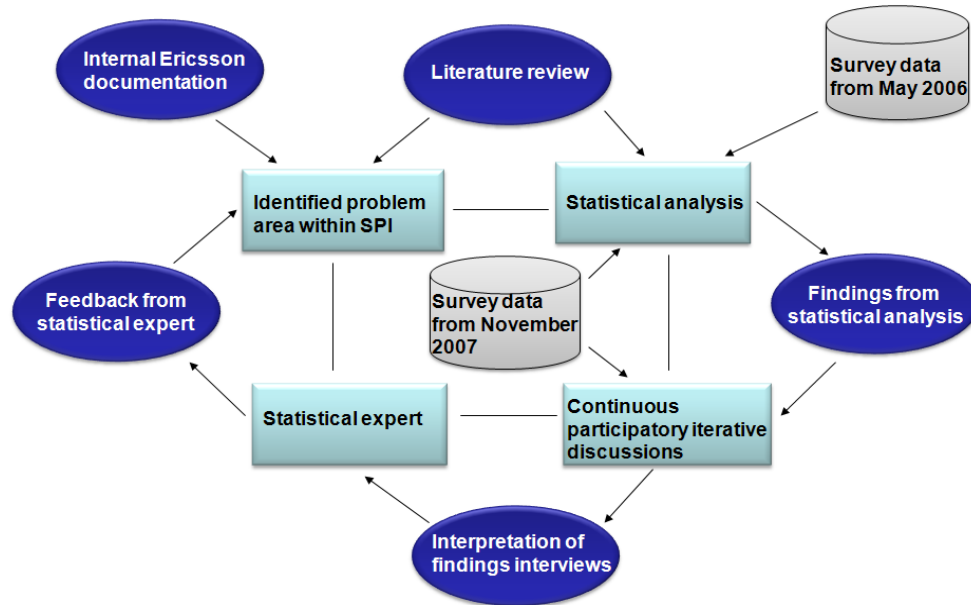


Fig. III: Conceptual design of research methodology

PARTICIPANTS

Involved in the research process, apart from the author, were a number of participants who contributed to the end results of this case study in various ways. The participants are listed in Table II.

Table II: Participants in the case study

#	Participant	Role	Contribution
1	Daniél Rocha	Author of the thesis	Literature review, conducting statistical analysis, writing thesis, organizing work procedure
2	Anna Börjesson Sandberg	Ericsson Supervisor	Feedback on statistical data, constant participant in discussions, expert on SPI in general and Ericsson employee process survey in particular. Feedback in writing process
3	Anders Baaz	Ericsson Supervisor	Feedback on statistical data, constant participant in discussions, expert on measurements in SPI in general and Ericsson employee process survey in particular
4	Stefan Stark	Ericsson external "statistical expert"	Feedback on results and findings from statistical analysis
5	The 6369 respondents of the employee process survey	Material that was statistically analyzed	Population that served as a base for the research and who's results were generalized at the end to answer the general research question
6	Björn Ohlsson	IT University Supervisor	Feedback and support during the thesis process
7	Lars Pareto	IT University Support	Contributed with literature recommendations and feedback on findings from statistical analysis

DATA SOURCES

The employee process survey from November 2007 served as the primary data source that was analyzed. The results from the statistical analysis of the real world measurement data were discussed with Ericsson SPI change agents and statistical experts to draw conclusions as to the validity and reliability of the results as well as the possibility to strengthen the quality of the results from SPI measurements within the organization. These discussions also formed a strong base for the generalization of the results to a larger population. The data collected and used throughout the case study is summarized in Table III.

Table III: Data sources used in the case study

#	Data sources	Comments
1	Results from employee processes survey	Data set from May 2006 and November 2007 that was used in the statistical analysis
2	Continuous participatory iterative discussions	Provided continuous feedback on statistical analysis results as well as insight into Ericsson procedures and culture
3	Literature	Provided knowledge on SPI as well statistical methods that was used
4	Internal documents at Ericsson	Provided data about general view of SPI activities at Ericsson, e.g. survey presentation material
5	Statistical Expert	Feedback on statistical output
6	Ericsson Change Agents	Provided information on SPI activities in general and SPI initiatives at Ericsson in specific. Participated in the continuous iterative interviews

THE ERICSSON MEASUREMENT CASE

New knowledge in the SPI measurement field is constantly needed in order for SPI activities carried out within organizations to be successful. As part of contributing to this body of knowledge Börjesson *et al.* (2007) performed research in the area of process innovations and improvement measurement mechanisms. Some of the findings from this research will be presented in the sections below to give the reader an insight into the history of the measurement mechanism at Ericsson. The layout of this chapter is as follows; first we present the background to the process model, following that the results from this research are presented.

BACKGROUND TO THE PROCESS MODEL

During a longitudinal action research project between 2001 to 2006 Börjesson *et al.* (2007) designed and implemented a measurement mechanism at a department at Ericsson. In the development of the process model four key ratios was developed; Processment, Process Commitment, Process Improvement and Process Learning. Each of these four key ratios served as a descriptor for a set of target areas within Ericsson's SPI work. The survey used to measure the process knowledge and use within Ericsson contained, at that time, ten questions. Each of these questions was in turn connected to one of the four key ratios as can be seen in Fig. IV below.

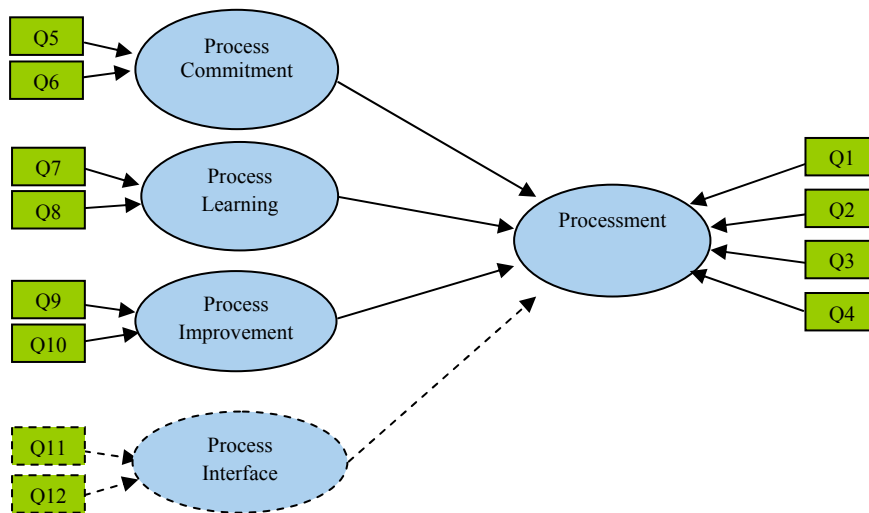


Fig. IV: View of the Process Model

This type of process model can support an organization in focusing their SPI measurements and analysis of the results. Continuing the work with the process model, Enskog (2006) conducted a research project at Ericsson where she analyzed the results of a process employee survey sent out in May 2006 and validated the process model with the data from this survey. With the validation of the model came a stronger belief in it. Yet the model was only tested for a one design unit within Ericsson R&D, with 1567 respondents. This leaves some uncertainties as to the validity of the process model.

In the process employee survey sent out in November 2007 two new questions were added, making a total of twelve questions. The two new questions formed a new key ratio, Process Interface. This key ratio is added to the process model, making a total of five key ratios (see Fig. IV). Additionally, the survey was this time answered by 6369 respondents. The larger number of respondents together with the new questions and key ratio allows for interesting analysis of the results and opens up for a good generalizability for several design units within Ericsson R&D. So there existed a need to perform sound statistical analysis on the latest results from the process employee survey.

RESULT AND ANALYSIS

In this section the statistical analysis of the employee process survey conducted in November 2007 are presented and analyzed. The section is divided into three major subsections; Survey Questions, Survey Key Ratios, and Partial Least Squares regression. In each of these subsections the results from the statistical methods performed are reported and analyzed as described in the chapter Theoretical Context and the section Basic Statistical Analysis.

SURVEY QUESTIONS

In this subsection the results from the statistical analysis methods performed on the answers to the questions from the survey are presented. We present the results according to the statistical analysis methods in the chapter Theoretical Context.

DISTRIBUTION ANALYSIS

We performed a normal distribution analysis on the answers to each question. The need to determine the distribution of the results was based on the fact that subsequent statistical methods required knowledge about the population's distribution. The results from the graphical methods, histogram with normal curve and normal Q-Q plot can be found in Fig. V and Fig. VI respectively. These

diagrams both report the results from analysis of question Q11. Results from the normal distribution analysis of the other eleven questions can be found in Appendix A.

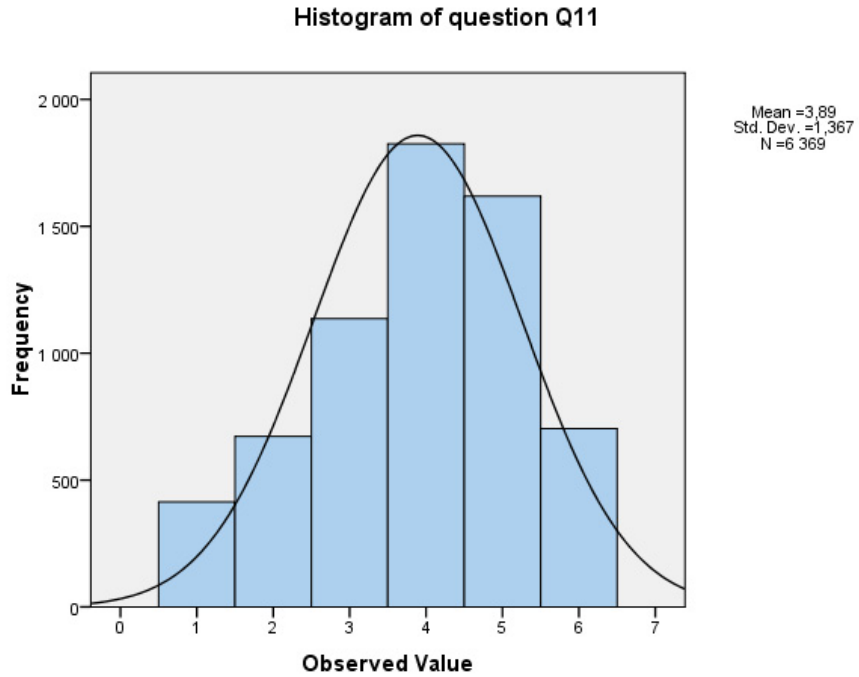


Fig. V: Histogram of question Q11 with normal distribution curve

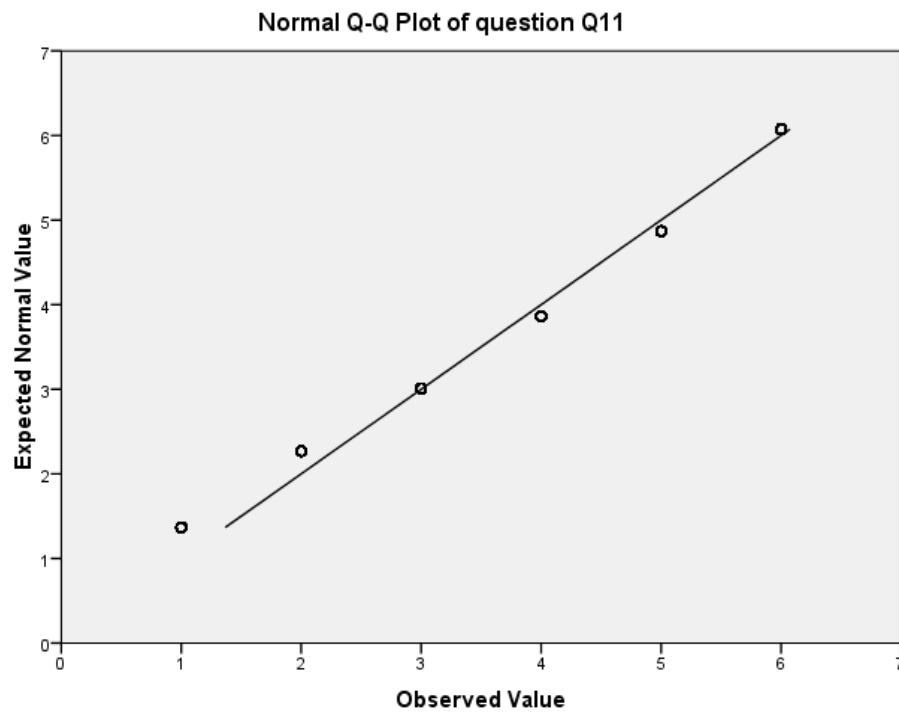


Fig. VI : Normal Q-Q Plot of question Q11

Through ocular inspection of the histograms presented we see that the results from answers to

FACTOR ANALYSIS

Besides the requirement of normality there existed a number of considerations to determine if the data was suitable for a factor analysis (see Factor Analysis). First of the population had to be large enough, however, with $n=6369$ respondents we were well above Tabachnick and Fidell's (2007) recommended threshold of $n>300$. The second consideration that we addressed concerned the inter-correlation among the questions, the R_S score. An inspection of the correlation matrix (see Table IV) reveals that the scores all range between 0,46 to 0,84 which indicates that a factor analysis could be considered.

To further assess if the data was suitable for a factor analysis Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's test for sphericity (see Factor Analysis) were also of interest to us. As reported in Table V we had a KMO index of 0,944, well exceeding the suggested threshold value for a sound factor analysis of 0,60 (Tabachnick and Fidell, 2007). Moreover, Bartlett's test of sphericity was significant $p<0,05$. Given these scores we concluded that the data was suitable for factor analysis.

Table V: Results from Kaiser-Meyer-Olkin Measure of Sampling Adequacy and Bartlett's Test of Sphericity

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0,944
Bartlett's Test of Sphericity	Significance	0,000

Factor analysis was used both to test the new data from the survey of 2007, against previous discoveries, (Enskog, 2006), as well as running it with the two additional questions. This allowed us to analyze any potential changes in the outcome when comparing the two results. For a comparison of the results to be possible question Q11 and Q12 is not included in Table VI, which reports results from factor analysis of the first ten questions from the survey of 2007. The colored cells in the table shows on what component each question converges, i.e. which component each question most relates to.

Table VI: Rotated Component Matrix – Results from factor analysis with the ten questions from 2007's survey

	Component			
	1	2	3	4
Q1	0,603	0,190	0,652	0,183
Q2	0,654	0,313	0,520	0,150
Q3	0,711	0,423	0,257	0,316
Q4	0,760	0,358	0,285	0,238
Q5	0,280	0,243	0,233	0,894
Q6	0,355	0,536	0,528	0,152
Q7	0,477	0,182	0,722	0,210
Q8	0,148	0,491	0,722	0,210
Q9	0,279	0,852	0,256	0,144
Q10	0,306	0,831	0,223	0,224

Extraction Method: Principal Component Analysis,

Rotation Method: Varimax with Kaiser Normalization,
Rotation converged in 9 iterations,
n = 6369.

When comparing the results with previous discoveries we could confirm the results presented by Enskog (2006), and since the sample size has increased to our n=6369, this indicates that the choice of the four hidden components (the key ratios) selected for the process model was valid. In order to analyze how the process model was affected by the two new questions (Q11 and Q12) added in the survey of 2007, we conducted a second factor analysis including the new questions. In Table VII below, the results from the factor analysis of all twelve questions from the survey of November 2007 can be found.

Table VII: Rotated Component Matrix – Results from factor analysis with the twelve questions from 2007's survey

	Component				
	1	2	3	4	5
Q1	0,838	0,225	0,208	0,182	0,204
Q2	0,758	0,307	0,318	0,148	0,158
Q3	0,593	0,363	0,503	0,303	0,003
Q4	0,656	0,303	0,489	0,228	-0,001
Q5	0,296	0,221	0,236	0,885	0,132
Q6	0,537	0,542	0,190	0,153	0,305
Q7	0,770	0,192	0,197	0,194	0,330
Q8	0,403	0,337	0,315	0,160	0,722
Q9	0,277	0,829	0,300	0,135	0,175
Q10	0,270	0,798	0,338	0,213	0,147
Q11	0,272	0,310	0,820	0,141	0,227
Q12	0,332	0,328	0,770	0,222	0,205

Extraction Method: Principal Component Analysis,
Rotation Method: Varimax with Kaiser Normalization,
Rotation converged in 7 iterations,
n = 6369.

With the inclusion of the two new questions, Q11 and Q12, we can see a slight change in the results compared to the results from the factor analysis performed with ten questions. This change can also be seen when comparing these results with the once presented in Enskog, (2006). Some of the questions from 2006's survey were changed slightly in the survey of 2007, for example, where the word feel was used before now the word believe is used instead. This could account for the slight displacement of questions Q7 and Q8 in the results. However, we can see that there are still some similarities with the previous results. Most importantly we can see indications that questions Q11 and Q12 loads on the same separate component which confirms the addition of the new key ratio Process Interface to the process model.

In the continuous participatory iterative discussions (see Fig. III) we discussed the implications of accepting the established process model even though some of the results from the factor analysis might suggest a different approach. Since factor analysis is not a definitive method and the interpretation of the results are largely up to the researcher (see Factor Analysis). We decided that the best approach was to not make any changes in the existing process model on account of to the slight displacement of some of the scores from the factor analysis. The cause behind this decision is the belief from our point of view that the process model is well established and therefore has a recognition value in the organization and more would be lost by abandoning it than keeping it. Moreover, if the process model should prove not to be of sufficient quality this would be revealed in the PLS

regression analysis that we performed. Should there have been any doubts as to the validity and reliability of the model, a step back from PLS regression to factor analysis would have been considered.

SURVEY KEY RATIOS

In this subsection the results from the statistical analysis methods performed on the key ratios will be presented and analyzed. First we present the results from the distribution analysis, secondly the results from the correlation analysis. Following that the results from the PLS regression analysis. The results are present according to the statistical analysis methods in the chapter Theoretical Context.

DISTRIBUTION ANALYSIS

We performed a normal distribution analysis on the survey key ratios for 208 organizational units. The results from the graphical methods, histogram with normal curve and normal Q-Q plot of the key ratio Processment, can be found in Fig. VII and Fig. VIII respectively.

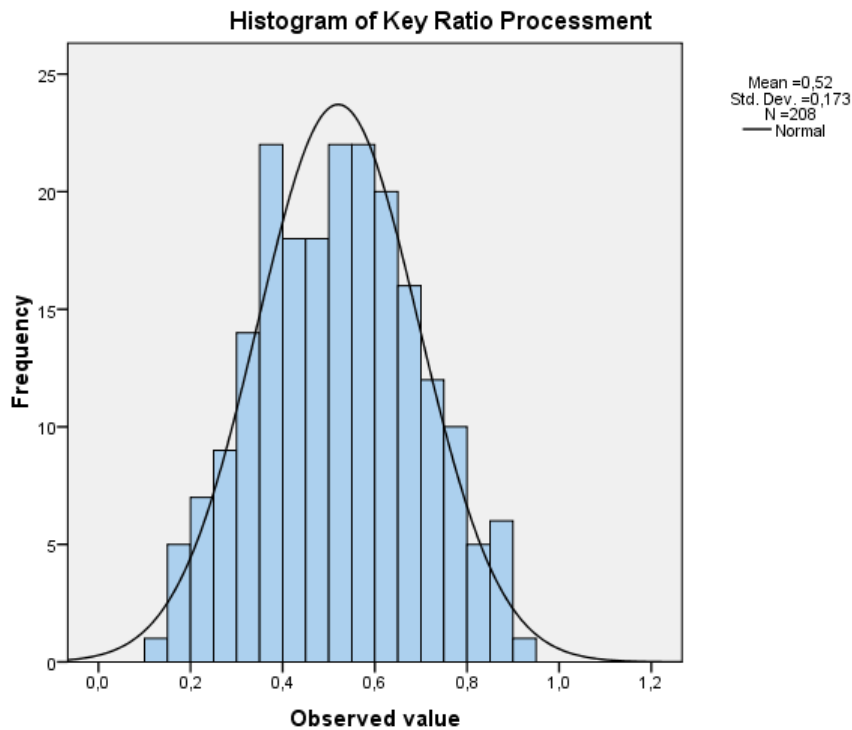


Fig. VII: Histogram of key ratio Processment with normal distribution curve

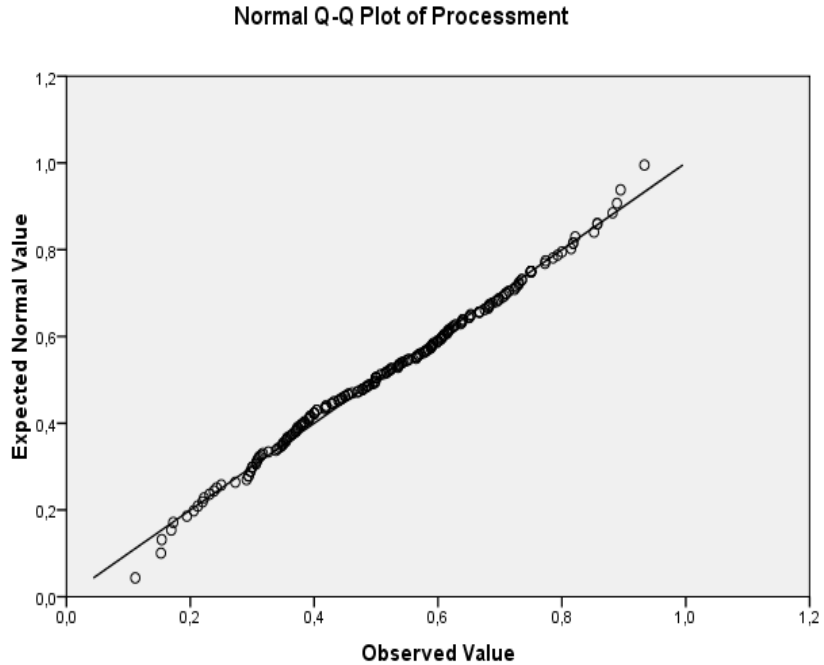


Fig. VIII: Normal Q-Q plot of Processment

As for the distribution of the key ratios we have seen that the key ratio Processment, in the graphical analysis of distribution (see Fig. VII), quite closely follows the expected normal distribution curve. We can also see that the result from the Normal Q-Q plot analysis of Processment (see Fig. VIII) further supports the assumption of normality for this key ratio. This is true also for the other key ratios. The results from the normal distribution analysis of these four can be found in Appendix B.

In order to further asses the assumed normality of the key ratios a One Sample Kolmogorov-Smirnov test was performed, the results from the K-S test reports a non-significant result ($p > 0,05$) indicating normality. We decided that the K-S test could be used to asses the distribution since our $n=208$ for the key ratios, which is not to large a sample for a sound K-S test (see Distribution Analysis). The results from this test can be found in Table VIII below.

Table VIII: Results from the One Sample Kolmogorov-Smirnov test for the five survey key ratios

	Processment	Process Commitment	Process Learning	Process Improvement	Process Interface
N	208	208	208	208	208
Kolmogorov-Smirnov Z	0,785	0,718	0,844	0,399	0,583
Asymp. Sig. (2-tailed)	0,568	0,681	0,474	0,997	0,886

Through the use of the graphical methods and the numerical K-S test to analyze the distribution of the key ratios we proven that we can assume normality in the distribution of them.

CORRELATION ANALYSIS

The analysis of the correlation between the survey key ratios shows that the Spearman's rho values all range between 0,64 to 0,81. A descriptive table of the results from the correlation analysis of the key ratios can be found in Table IX.

Table IX: Spearman's rank order correlation analysis of key ratios

	Processment	Process Commitment	Process Learning	Process Improvement	Process Interface
Processment	1,00	0,77	0,81	0,69	0,75
Process Commitment		1,00	0,70	0,70	0,68
Process Learning			1,00	0,64	0,65
Process Improvement				1,00	0,70
Process Interface					1,00

These results indicate that high scores on one key ratio usually mean high scores on the other positively related key ratios. It also shows us that all of the key ratios target the same phenomenon in slightly various ways. This means that should the SPI organization at Ericsson target one improvement area, represented here by one key ratio, with SPI activities it is likely that the other areas will increase as well. The strong correlation results of the key ratios provided us with a first indication that they all belonged to the process model. However, to properly assess the process model a more thorough analysis was required, we therefore turned to PLS regression.

PARTIAL LEAST SQUARES REGRESSION ANALYSIS

With the confirmation from the factor analysis and with the addition of the new key ratio to the process model we turned to PLS regression analysis. This analysis would prove whether the process model was indeed valid and reliable and we could compare the results from this analysis with the once presented by Enskog (2006). First of we will present and discuss the results from the analysis of the measurement model, followed by the structural model.

SmartPLS was used to test the process model with the data from the process employee survey of 2007. With the techniques available in SmartPLS we could analyze the latent constructs, the key ratios, from a set of indicators, the survey questions (Chin, 1996). The PLS regression analysis was conducted to test the validity and reliability of both the measurement model and the structural model, results of which both are presented in the sections below. The scores for loading, average variance extracted, composite reliability, communality, cronbach alpha, path coefficients and their significance level, as well as the coefficient of determination were all calculated in this process.

MEASUREMENT MODEL

To assess the robustness and reliability of the measurement model we analyzed the scores for; average variance extracted, composite reliability, cronbach alpha and communality. The results from these coefficients are reported in Table X.

Table X: Results of coefficients for assessing the reliability of the measurement model

	Average Variance Extracted	Composite Reliability	Cronbach Alpha	Communality
Processment	0,81	0,94	0,92	0,75
Process Commitment	0,75	0,86	0,67	0,92
Process Learning	0,82	0,90	0,78	0,93
Process Improvement	0,92	0,96	0,91	0,82
Process Interface	0,93	0,96	0,92	0,81

The results from this reliability test of the measurement model showed us that all the scores were well above or just around the suggested thresholds. AVE range from 0,75 to 0,93, well exceeding the suggested threshold of 0,50 for all constructs. The score for Composite Reliability range from 0,86 to 0,96, also exceeding the recommended threshold value of 0,70. Inspection of Cronbach's Alpha scores reveals that the scores range from 0,67 to 0,92 showing well enough scores. Finally, scores for Communality range from 0,75 to 0,93 exceeding threshold value of 0,50 for all key ratios. This shows us that the latent constructs (key ratios) are well explained by their corresponding indicators (the questions) and it indicates robustness and reliability in the model.

CONVERGENT AND DISCRIMINANT VALIDITY

We have seen that the correlations between the questions were strong (see Table IV) as well as the correlation among the key ratios (see Table IX). The next step was to test the measurement model for validity through analysis of convergent and discriminant validity. Convergent validity assess whether each proposed indicator to construct relationship was valid (see Table XI).

Table XI: Test of convergent validity. Factor loadings (in Bold) and cross loadings

	Processment	Process Commitment	Process Learning	Process Improvement	Process Interface
Q1	0,89	0,69	0,80	0,59	0,61
Q2	0,91	0,71	0,76	0,63	0,66
Q3	0,90	0,73	0,71	0,69	0,75
Q4	0,90	0,71	0,71	0,64	0,72
Q5	0,63	0,85	0,56	0,53	0,57
Q6	0,74	0,89	0,70	0,68	0,63
Q7	0,79	0,66	0,92	0,58	0,60
Q8	0,71	0,67	0,90	0,66	0,65
Q9	0,67	0,67	0,65	0,96	0,68
Q10	0,69	0,68	0,65	0,96	0,70
Q11	0,71	0,64	0,64	0,68	0,96
Q12	0,76	0,69	0,68	0,71	0,96

In our analysis of convergent validity, table above, we have seen that all of the indicators (questions) load on their own latent constructs (key ratios) with a value that exceeds 0,70, which is the threshold recommended by (Trochim, 2006). In addition, each question load higher on their own latent constructs than on the others. The analysis of convergent validity shown that each indicator is well correlated with the construct it is connected to.

Given the strong correlation of key ratios (see Table IX) it is important to asses that the constructs are not redundant, i.e. that two constructs in the end measures the same phenomenon. Through our analysis of discriminant validity we examined the degree to which the constructs diverged from each other (see Table XII).

Table XII: Test of discriminant validity. Inter-correlations of latent constructs

	Processment	Process Commitment	Process Learning	Process Improvement	Process Interface
Processment	0,899	0,791	0,829	0,709	0,763
Process Commitment		0,868	0,732	0,705	0,689
Process Learning			0,906	0,677	0,688
Process Improvement				0,958	0,720
Process Interface					0,962

We can see that the square roots of the AVE's (the bold figures) for each item are greater than their correlation with the other constructs (see Table XII), which indicates that the key ratios do in fact measure different concepts (see Partial Least Squares (PLS) Regression Analysis). This, in turn, indicates validity of the measurement model (Heeler and Ray, 1972).

We conclude that the measurement model is robust and reliable as was suggested by Enskog (2006). Moreover the validity of the measurement model is strengthened through these results as the number of respondents of the survey has increased dramatically since the previous results were presented.

STRUCTURAL MODEL

To assess the relationships of the constructs in the structural model we used the coefficient of determination, R^2 as well as the path coefficients and their corresponding significance score, which were retrieved from a t-table (see The Structural Model). We calculated the results of R^2 for the dependant construct Processment as well as path coefficients between the independent constructs, Process Commitment, Process Learning, Process Improvement and Process Interface and the dependant construct Processment. The results from the analysis of the structural model can be found in Fig. IX.

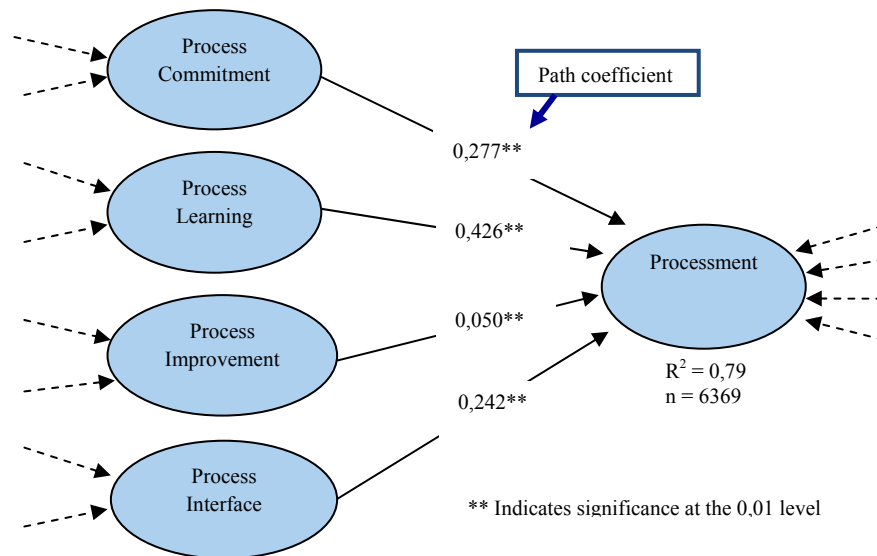


Fig. IX: Structural model with R^2 and path coefficient scores (depicts the structural model, leaving out the indicators that are part of the measurement model).

Through PLS regression analysis of the structural model we can see that the loadings for the two new questions Q11 and Q12 exceeds the recommended threshold of 0,80 (Chin, 1995) indicating a strong relationship with the new key ratio Process Interface (see Table XI). We can also see that Process Interface is a strong determinant of Processment with a path coefficient of 0,242 indicating significance at the 0,01 level (see Fig. IX). This result indicates that key ratio Process Interface does in fact belong in the process model.

Moreover, the results from the analysis of the structural model, without including the new key ratio process interface, showed us that the process model is strong, indicating that the constructs; Process Commitment, Process Learning, and Process Improvement are good determinants of Processment. The independent constructs explain 77% of the variance of Processment. This confirms the results presented by Enskog (2006), which reported a value of 76%. Furthermore, the number of respondents, n , has increased since the survey of May 2006, which enhances the validity of the model. When the results had been validated against previous results (Enskog, 2006) we added the two new questions and the new key ratio to the model. With the addition of the new questions and the new key ratio to the process model the four constructs; Process Commitment, Process Learning, Process Improvement, and Process Interface are good determinants of Processment. These constructs account for 79% of the variance of the dependant construct Processment (see Fig. IX). We conclude that there is consistency in the structural model and that the model is reliable.

DISCUSSION

In this chapter we discuss and reflect on the implications that this research could have on Ericsson based on what is reported in the chapter Result. The discussion and reflections relates back to the theories which are presented in the chapter Theoretical Context. The intention of the discussion is to present how we analyzed the outcome of the statistical methods performed and the results generated by them. Moreover, we intend to show how we fulfilled the general purpose of this research through the results that was reach in this study. The purpose of this study was:

Assuring quality of results from SPI measurements through statistical analysis.

The layout of this chapter is as follows; first we present our Reflections on the Statistical Analysis of the survey results and secondly we discuss the Implications for Practice that this study generated.

REFLECTIONS ON THE STATISTICAL ANALYSIS

In this study we have applied a set of statistical methods (see Basic Statistical Analysis) on SPI measurements, in this case in the form of survey results. In this chapter we reflect on the outcome of the statistical analysis and the potential gain of using such methods.

In our case we applied correlation analysis to the answers to the questions which was a vital first step that ensured that each of the questions actually had a relationship amongst them (see Table IV). As described before by using this method we could assume that none of the questions were out of place in the survey. Furthermore, as the number of respondents in the survey of November 2007 were 6369, the generalizability to a several other design units within Ericsson is strong. We believe that in future surveys, which might include even larger parts of the organization, the correlations between the questions should still exist as long as the wording of the questions does not change. Therefore the need to perform such an analysis might decrease for Ericsson. However, as future surveys might target different populations within the organization than this one has, the potential changes in the results call for correlation analysis to be performed. Furthermore, we see the need to ensure strong relationships between the answers to the questions in order to create validity which is needed in succeeding analysis methods. This leads us to suggest that a correlation analysis should still be performed in future undertakings.

Another useful statistical method that was performed during this study was the factor analysis (see Factor Analysis). The factor analysis showed us, with minor deviations, that the selection of the five key ratios; Processment, Process Commitment, Process Improvement, Process Learning, and Process Interface, all seem to be appropriate components for describing the questions (see Table VII). With the statistical finding that these five key ratios do in fact all have their place in the process model, the factor analysis in future statistical analysis of survey results would not prove as important as the one performed in this study, so long as no new questions or key ratios are added. However, performing a factor analysis provides a good indication as to how each question in the survey loads on what component. Furthermore, since there were some displacements in the factor analysis results (see Factor Analysis), it could be interesting for Ericsson to perform a new factor analysis with future survey data in order to monitor this displacement. On the other hand we feel, given that factor analysis is not a definitive method for extracting valid components, the results from such an analysis is not to be considered the final verdict.

We also wanted to assess the process model that is used within Ericsson to describe the results of the process survey. So we turned to PLS regression analysis (see Partial Least Squares (PLS) Regression Analysis). We could see that this technique has great potential, not only as a powerful statistical method for analyzing relationships and significance of survey results, but also as a method that outputs easy accessible results that, once interpreted by change agents, can be presented and understood to large parts of an organization. We believe that through the use of PLS regression it is possible that the perceived usefulness of the SPI measurements becomes apparent to a larger population, which in turn increases the chances of them being accepted (Davis, 1989). The perceived usefulness is however something that should be measured separately for any valid conclusions to be drawn from it. This will be discussed further in the section Implications for Practice

We feel that the most important statistical method used to analyze the survey data in this study was in fact the PLS regression analysis. As mentioned in Partial Least Squares (PLS) Regression Analysis, the method is well suited for ordinal survey data as it does not require an assumption of normality. The method is however somewhat dependant on a set of statistical methods that we suggest should be performed first to ensure that the results from the PLS regression analysis are reliable. The steps reflected upon in this section are summarized in Fig. X.

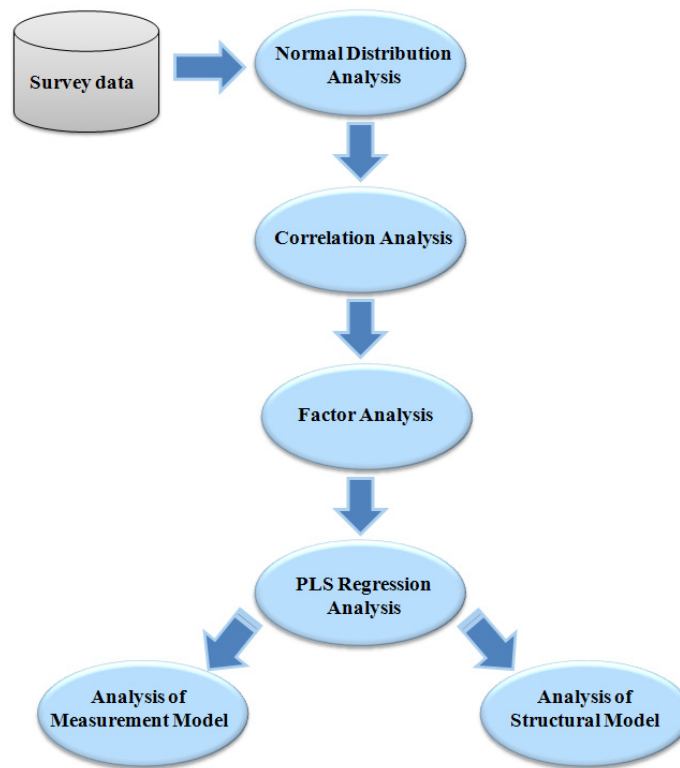


Fig. X: Suggested steps when performing Statistical Analysis on SPI Measurements

As for the assessment of the measurement and structural model in PLS regression analysis, there are a number of different software's available on the market, but we feel that SmartPLS well enough does the job. We believe that by using a program such as SmartPLS facilitates the process of conducting such an analysis. As of today, the software SmartPLS is free of charge. It incorporates the most important functions in PLS regression analysis. There are a number of steps in assessing the measurement and structural model. Most of the values needed to asses both models are received when performing PLS regression analysis in SmartPLS. We believe that there are a number of important statistical scores that needs to be considered when evaluating the models and these are presented in the section Partial Least Squares (PLS) Regression Analysis.

We feel that potential benefits from using statistical analysis on SPI measurements for an organization, in this case Ericsson; is a vital part in an ongoing SPI initiative. Bryman and Cramer (1994) stated the importance of awareness of quantitative data analysis in order to identify incorrect conclusions and manipulation of SPI measurement results. We have seen that the analysis performed resulted in a strong statistical validity and reliability of the measurements. Moreover the need to asses the assessments has been stated by Argyris (1982). We believe that this should be done in order to create confidence in the SPI organization among the employees and we suggest the use of the statistical methods presented in Fig. I as a good way of achieving this. However, no conclusions as to how these results are accepted within the organization can be drawn at this stage. The need for more studies where methods such as, interviews with employees, is needed to confirm the impact on the beliefs of the measurements before and after the application of statistical analysis.

This set of statistical methods can be reused, which facilitates continuous systematic monitoring in future SPI initiatives, the benefits of which have been stated by Zahran (1997). How these statistical methods fit in to the SPI activities at Ericsson and how they can be used at Ericsson are further discussed in the section Implications for Practice.

In future undertakings we see the need to perform an analysis on the two new questions and the resulting key ratio in the survey of November 2007. If we look closer at the wording of the questions,

we can see that they focus on input output; this indicates an interesting interaction between units. Where question Q11 focuses on the perceived quality of the input one receives from other units question Q12 focuses on the perceived quality of the output that one generates. We feel that the interaction on unit level and the correlation of perceived quality between one units output and the subsequent input that it becomes for the receiving unit would be an interesting area to focus on. We believe that such an analysis could potentially prove that there in fact exist inconsistencies in the two results in between units. With the results from such an analysis the SPI change agents at Ericsson would have a good insight into the different units and their own perceived strengths and weaknesses in the material produced as well as the other interacting units' views on the same matter.

We conclude that the set of statistical methods and tools suggested in this study can be used to assure the quality of the results from SPI measurements.

We conclude that the statistical analysis of the process employee survey investigated in this thesis shows strong consistency and validity of both survey questions and key ratios. We also conclude that we can strengthen the statistical validity in SPI measurements through the use of basic statistical analysis

IMPLICATIONS FOR PRACTICE

There are a number of ways that the set of statistical methods presented in this paper can be beneficially used within Ericsson. We can see that the need to incorporate statistical analysis of SPI measurements in order to strengthen the statistical validity and quality of said measurements.

The first practical contribution of this study is the potential for Ericsson to educate their SPI change agents to use the statistical methods suggested in this paper in future SPI initiatives. This will ensure the validity of the results from SPI measurements such as the survey that Ericsson conducts continually. We believe that educating Ericsson's change agents in the suggested statistical methods should support the SPI organization in future SPI initiatives.

The second practical contribution for Ericsson that we present in this study is the proposed four phase iterative SPI measurement analysis process, presented in Fig. XI. This model should be possible to integrate in the continuous iterative SPI activities currently conducted at Ericsson.

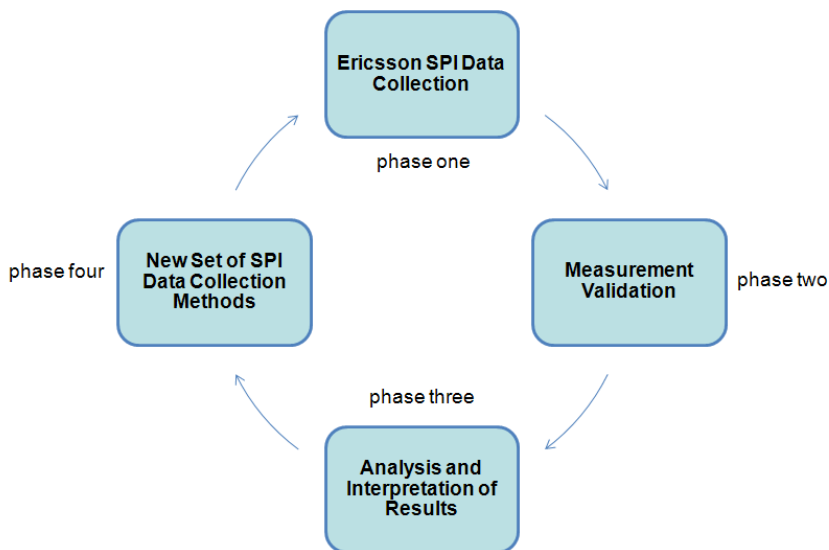


Fig. XI: Conceptual model of the suggested four phase iterative SPI measurements process

The first phase in the model, Ericsson SPI Data Collection, is the one already present, where Ericsson's change agents' collects SPI measurement data in the form of a survey and the process used to define this. We suggest, based on the findings in this document, the second phase of our model, Measurement Validation in this case performed by the author in collaboration with Ericsson. In this phase we feel that Ericsson would benefit from performing the statistical analysis steps presented in Fig. X and reported in this study. In the third phase, Analysis and Interpretation of Results, the results from the validation in phase two should be analyzed for validity and quality as suggested in Fig. X. In this phase the interpretation and analysis should also be performed together with the initial conclusions received in the first phase by Ericsson's change agents', which we believe should result in an additional set of conclusions. The final phase of our suggested model is the, New Set of SPI Data Collection Methods, where the conclusions drawn in phase three affect the data collection in the following phase one with new methods for SPI data collection as the process comes full cycle. This way of working with measurements as an integrated improvement model in the continuous SPI process is similar to the suggested activities of the higher levels of CMM (Paulk *et al.* 1995).

Thirdly, we can see three ways for Ericsson to evaluate the potential impact of performing statistical analysis on the SPI measurements among the employees:

- A. The idea of keeping one control group in each SPI measurement. This group would only perform the first phase in the suggested model (Fig. XI), Ericsson AB SPI Data Collection, while the other group performs all four phases in the model. The difference in the results of these two groups could then be used to asses if there are any significant differences in the perceived benefits among the two groups.
- B. The second suggestion is a directed study, where the participants could be subjected to deeper qualitative interviews or questionnaires. The data gathered from this could then be used to investigate the impact that the statistical analysis has on the acceptance of the results among the employees.
- C. We believe that there is also a third option. A mixed methods approach where methods and techniques presented in suggestions A and B, above, are both represented. This method is however somewhat more time consuming, so the need for the organization to start evaluating the potential ROI in such an initiative is suggested.

This study did not set out to formulate any implications for theory. However, at the conclusion of the study we can see the potential of such implications being presented. Should the four phases in the suggested iterative SPI measurement process (see Fig. XI) be continuously followed by Ericsson and if there were other similar studies conducted and analyzed at other organizations, a set of theoretical implications could be derived. This off course requires analysis of how the employees' beliefs in SPI initiatives are affected by the statistical analysis through deeper qualitative studies, as suggested above. This was however out of the scope of this study and therefore not further explored herein.

We conclude that the four phase iterative SPI measurements process model presented could be used to assure quality in SPI measurements and improve future SPI measurements.

CONCLUSIONS

In this thesis we investigate how we can strengthen the statistical validity of SPI employee survey measurements at the Ericsson telecom company with help of statistical analysis. We have used basic statistical analysis methods such as distribution analysis, correlation analysis, factors analysis and partial least squares regression analysis on a process employee survey sample consisting of 12 answers

from 6369 individuals. We conclude that the statistical analysis of the process employee survey investigated in this thesis shows strong consistency and validity of both survey questions and key ratios, and the process model created from the two shows a strong robustness and reliability. We also conclude that the set of statistical methods and tools suggested in this study can be used to assure the quality of the results from SPI measurements. Finally we conclude that the four phase iterative SPI measurements process model presented could be used to assure quality in SPI measurements and improve future SPI measurements.

ACKNOWLEDGEMENT

The author would like to extend a warm thanks to Anna Börjesson Sandberg and Anders Baaz at Ericsson AB for their untiring support and encouragement during this thesis process. Thanks also to Björn Olsson at the IT University for pushing me to solve many of the issues encountered during the writing of this paper. Appreciation also to William Sullivan at the IT University and Stefan Stark at Guide Consulting for taking time out of their busy schedule to answer questions and providing insights into the vast field that is statistics. Thanks.

REFERENCES

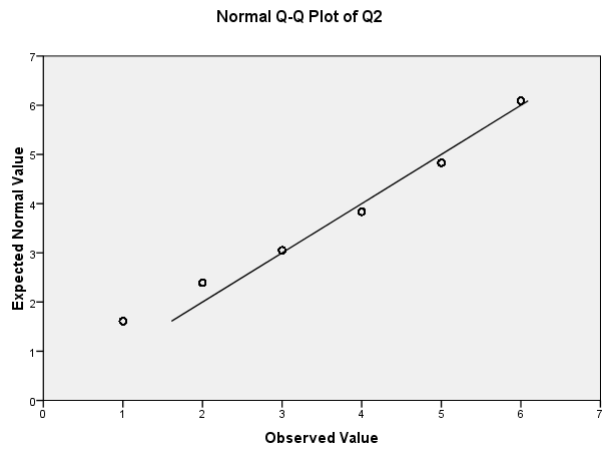
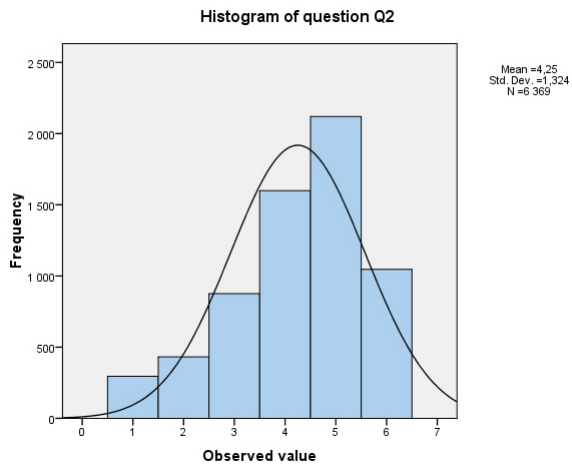
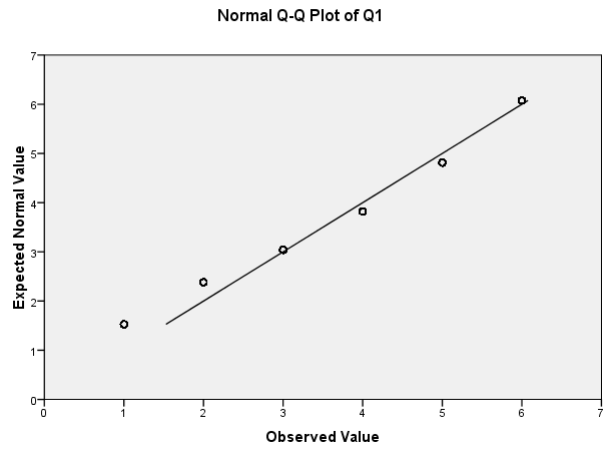
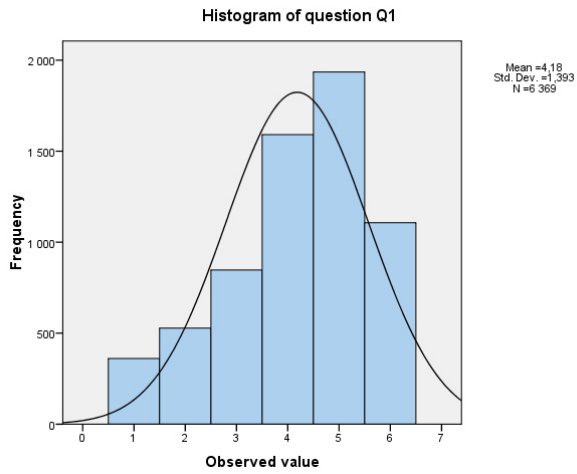
- Argyris, C., (1982) "The Executive Mind and Double-Loop Learning". *Organizational Dynamics*, Autumn 1982, AWCOP Periodicals Division, American Management Associations.
- Bartlett, M.S. (1954). "A note on the multiplying factors for various chi square approximations." *Journal of the Royal Statistical Society*, 16 (Series B), 296-8.
- Bernstein, S.N. (1945). "On the work of P.L.Chebyshev in Probability Theory". *The Scientific Legacy of P. L. Chebyshev. First Part: Mathematics* Edited by S. N. Bernstein. *Academiya Nauk SSSR, Moscow-Leningrad.*
- Brancheau, J., Janz, B., and Wetherbe, J., (1996). "Key Issues in Information Systems Management: 1994-1995 SIM Delphi Results", *MIS Quarterly*, Vol. 20, No 2, 1996, pp. 225-242.
- Bryman, A. and Cramer D., (1994). "Quantitative Data Analysis for Social Scientists". Department of Social Sciences, Loughborough University of Technology.
- Börjesson, A. and Mathiassen. L., (2003). "Making SPI Happen: The IDEAL Distribution of Effort", IT University, Gothenburg, Sweden, Ericsson AB, St Sigfridsgatan 89, SE-412 66 Gothenburg, Sweden.
- Börjesson, A. and Mathiassen, L., (2004). "Successful Process Implementation". *IEEE Software*, Vol. 21, No. 4, pp. 36-44.
- Börjesson. A., (2006). "Simple indicators for tracking software process improvement progress", Ericsson AB and IT University of Gothenburg, Lindholmospiren 11, 417 56 Gothenburg, Sweden.
- Börjesson A., Baaz A., Pries-Heje J., and Timmerås M., (2007). "Measuring Process Innovations and Improvements", In *IFIP International Federation for Information Processing, Organizational Dynamics of Technology-Based Innovation: Diversifying the Research Agenda*, (Boston: Springer).
- Chin, W. W., (1995). "Partial Least Squares is to Lisrel as Principal Components Analysis is to Common Factor Analysis". *Technology Studies*, 2, Faculty of Management, University of Calgary, Calgary, Alberta T2N 1N4, Canada.

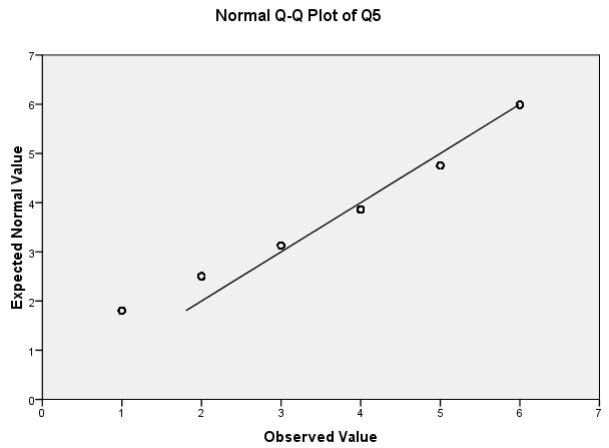
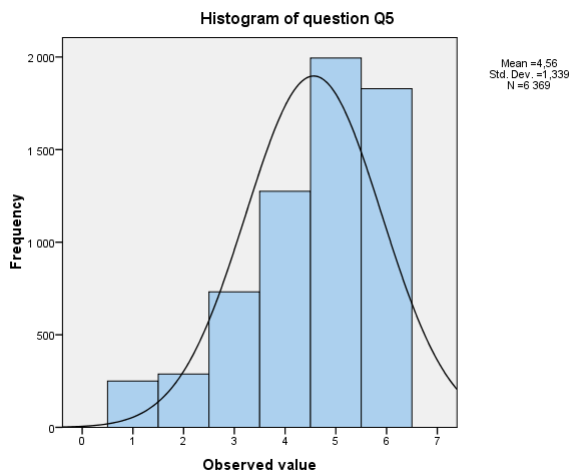
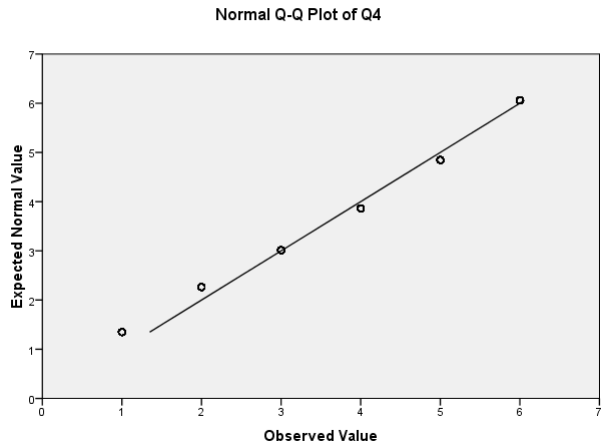
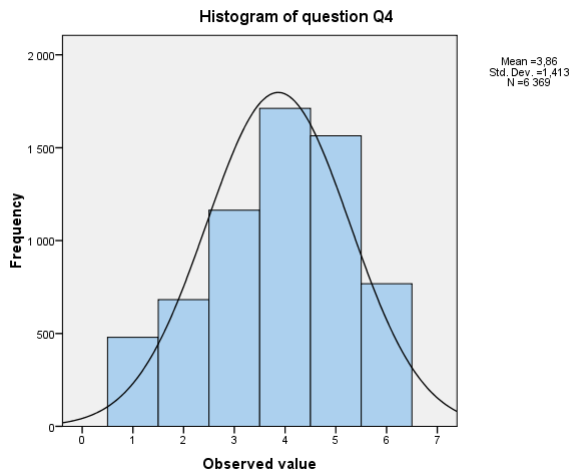
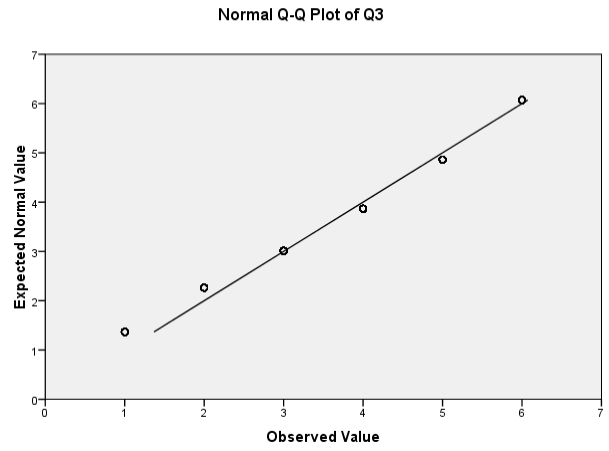
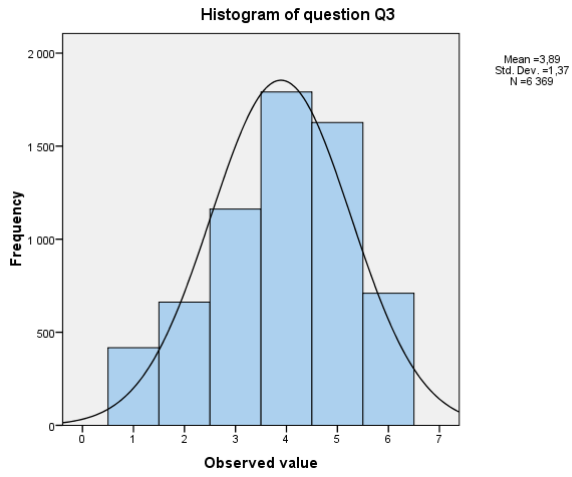
- Chin, W. W., Marcolin, B. L. and Newsted, P. R., (1996). "A Partial Least Squares Latent Variable Modeling Approach for Measuring Interaction Effects: Results from a Monte Carlo Simulation Study and Voice Mail Emotion/Adoption Study", in Proceedings of the International Conference on Information Systems, DeGross, J.I., Srinivasan, A. and Jarvenpaa, S. (editors), Cleveland, OH, 1996, pp. 21-41.
- Chin, W. W., (1998). "The Partial Least Squares Approach for Structural Equation Modeling", in Modern methods for business research (Marcoulides, G. A. Ed.), Lawrence Erlbaum Associates, Hillsdale.
- Cohen, J.W., (1988). "Statistical power analysis for the behavioral sciences" (2nd edn). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Davis, F. D., (1989). "Perceived usefulness, perceived ease of use, and user acceptance of information technology", MIS Quarterly, Vol 13, 1989, pp. 319–340.
- DeVellis, R.F., (2003). "Scale development: Theory and application (2nd edn). Thousand Oaks, California. Sage.
- Dickson, G., Leitheser, R., Wetherbe, J., and Nechis, M., (1984). "Key Information Systems Issues for the 1980's", MIS Quarterly Vol. 8, No 3, 1984, pp. 135-139.
- Dillon, W.R. and Goldstein M., (1984). "Multivariate Analysis: Methods and Applications", New York. Wiley.
- Dybå, T., (2000). "An Instrument for Measuring the Key Factors of Success in Software Process Improvement", Empirical Software Engineering Vol. 5, 2000, pp. 357-390.
- Enskog, U., (2006). "Process Awareness and Use at Ericsson AB," B.Sc. Thesis, IT University of Göteborg, Sweden.
- Fornell, C. and Larcker, D., (1981). "Evaluating structural equation models with unobservable variables and measurement error". Journal of Marketing Research. Vol. 18, n. 1, Feb. 1981, p. 39-50.]
- Grady, R.B., (1992). "Practical Software Metrics for Project Management and Process Improvement", Upper Saddle River, New Jersey, Prentice Hall.
- Grady, R. B., (1997). "Successful Software Process Improvement", Upper Saddle River, NJ: Prentice Hall.
- Gravetter, F.J and Wallanau, L.B., (2004). "Statistics for the behavioral sciences" (6th edn). Belmont, CA, Wadsworth.
- Hartog, C. and Herbert, M. (1985). "Opinion Survey of MIS Managers: Key Issues", MIS Quarterly Vol. 10, No 4, 1986, pp. 351-361.
- Heeler, R.M. and Ray, M.L., (1972). "Measure Validation in Marketing," Journal of Marketing Research, 9 (November), 361-70.
- Herbsleb, J., Carleton, A., Rozum, J., Siegel, J., and Zubrow, D., (1994). "Benefits of CMM-Based Software Process Improvement: Initial Results," Technical Report, CMU/SEI-94-TR-13, Carnegie Mellon University, Software Engineering Institute, Pittsburgh, PA.
- Humphrey, W., (1989). "Managing the Software Process", Addison Wesley Professional

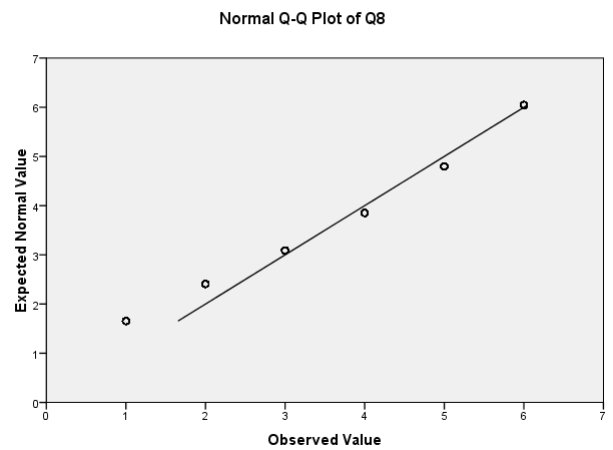
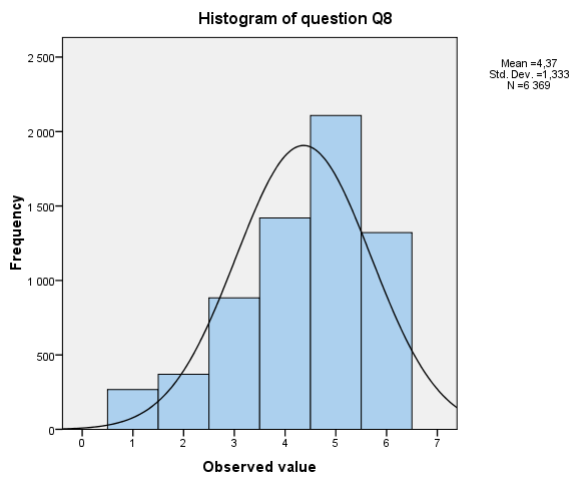
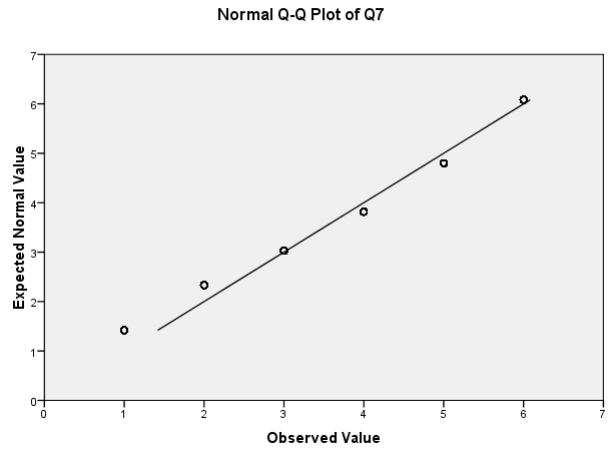
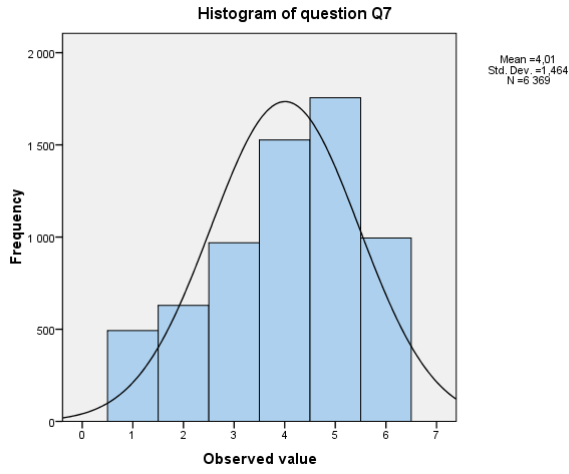
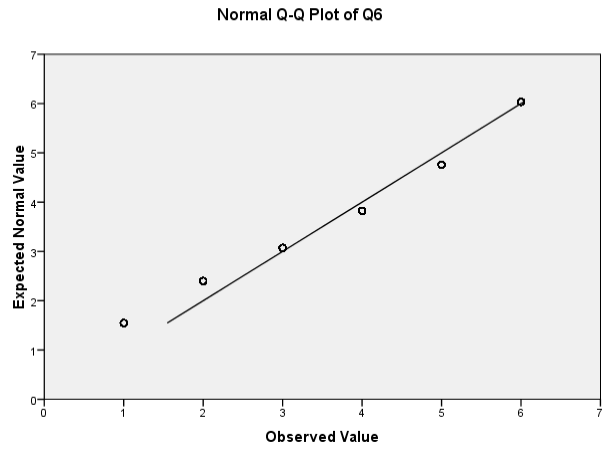
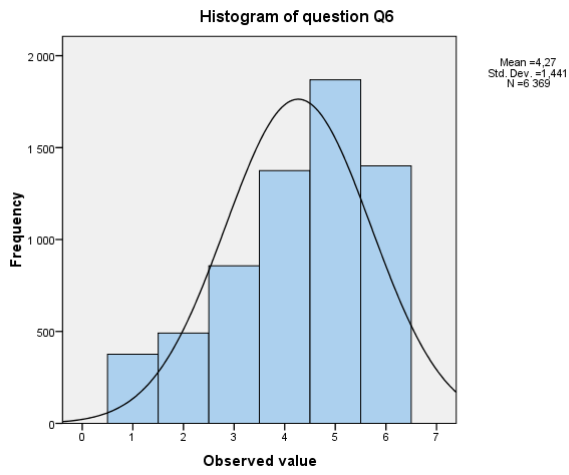
- Iversen, J. and Ngwenyama, O. (2006). "Problems in measuring effectiveness in software process improvement: A longitudinal study of organizational change at Danske Data". University of Wisconsin Oshkosh, 800 Algoma Blvd., Oshkosh, WI 54901-8675, USA, Ryerson University, Toronto, Canada Virginia Commonwealth University, Richmond, VA 23284-4000, USA.
- Kaiser, H., (1970). "A second generation Little Jiffy". *Psychometrika*, 35, 401-15.
- Kaiser, H., (1974). "An index of factorial simplicity". *Psychometrika*, 39, 31-6.
- Marasinghe, M., Meeker, W., Cook, D. and Shin, T.S., (August, 1994). "Using graphics and simulation to teach statistical concepts", Paper presented at the Annual meeting of the American Statistician Association, Toronto, Canada.
- McFeeley, B., (1996). IDEAL. "A User's Guide for Software Process Improvement", The Software Engineering Institute, Carnegie Mellon University, Pittsburgh, Handbook CMU/SEI-96-HB-001.
- Niederman, F., Brancheau, J., and Wetherbe, J., (1991). "Information System Management Issues for the 1990s", *MIS Quarterly* Vol. 15, No 4, 1991, pp. 475-500.
- Paulk, M.C., Weber, C.V., Curtis, B., and Chrissis, M.B., (1995). "The Capability Maturity Model: Guidelines for Improving the Software Process". Reading, Mass., Addison-Wesley Pub. Co.
- Pólya, G., (1920). "Ueber den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung und das Momentenproblem", *Math. Zeit.* Vol. 8.
- SEMA, (2002). "Process Maturity Profile of the Software Community", Software Engineering Institute, Carnegie-Mellon University.
- SmartPLS 2.0, M3 Release, (May, 2008). University of Hamburg School of Business Institute for Operations Management and Organizations. Von-Melle-Park 5, 20146 Hamburg, Germany. [Online]. Available: <http://www.smartpls.de/>
- SPSS Inc. (May, 2008). SPSS Inc. Headquarters, 233 S. Wacker Drive, 11th floor Chicago, Illinois 60606. [Online]. Available: <http://www.spss.com/>
- Tabachnick, B.G. and Fidell, L.S., (2007). "Using Multivariate Statistics" (5th edn). Pearson Education, Boston.
- Trochim, W.M., (Oct, 2006). "Research Methods Knowledge Base", [Online], Available: <http://www.socialresearchmethods.net/kb/contents.htm>
- Weinberg, G.M., (1993). "Quality Software Management Volume II: First-Order Measurement". Dorset House Publishing, New York, USA.
- Weinburg, G.M., (1997). "Quality Software Management, Volume IV: Anticipating Change". Dorset House Publishing, New York.
- Yoo, Y. and Alavi, M., (2001). "Media and Group Cohesion: Relative Influences on Social Presence, Task Participation, and Group Consensus", *MIS Quarterly* Vol. 25, No 3, 2001, 371-390.
- Zahran, S., (1997). "Software Process Improvement: Practical Guidelines for Business Success". Addison-Wesley.
- Zmud, R.W., (1984). "An Examination of "Push-Pull" Theory Applied to Process Innovation in Knowledge Work". *Management Science*, Vol. 30, No. 6 (Jun.), pp. 727-738.

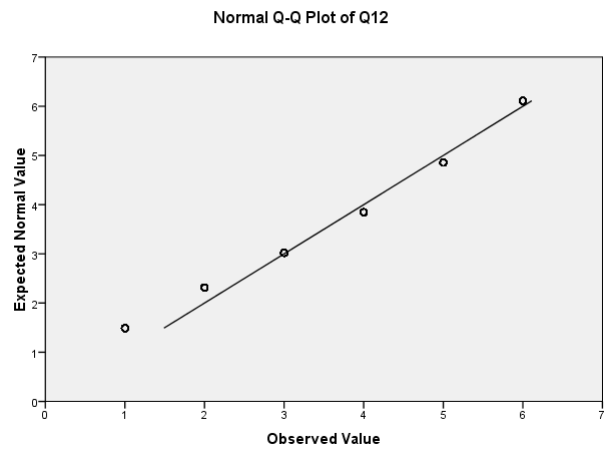
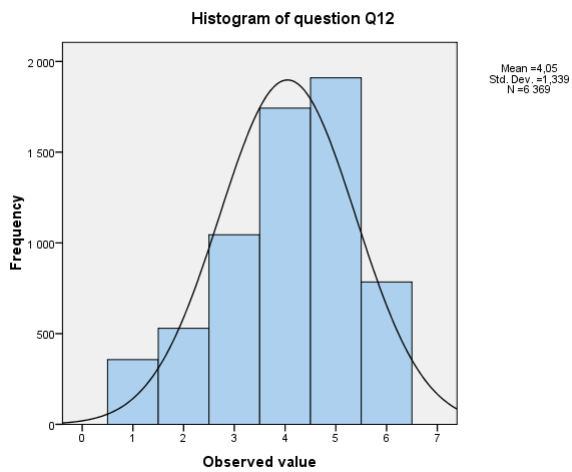
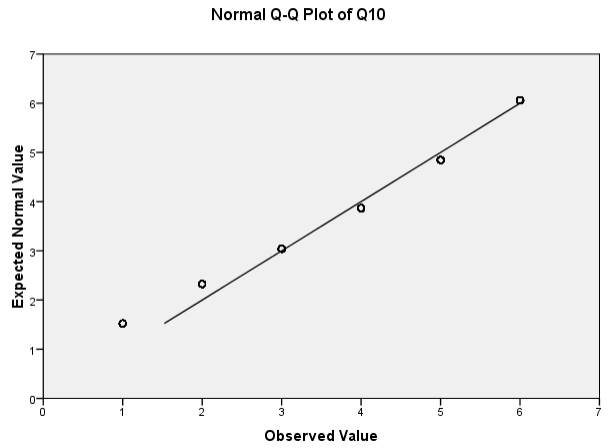
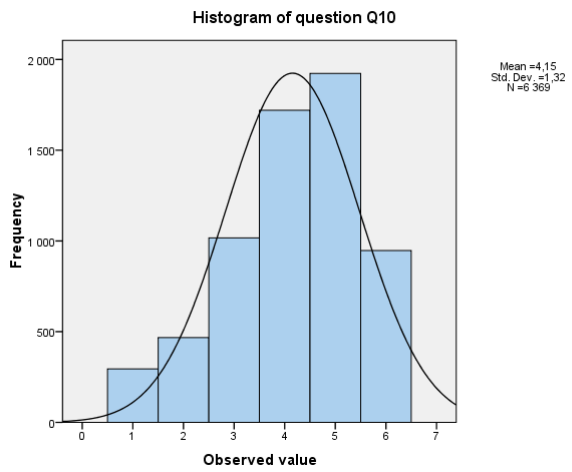
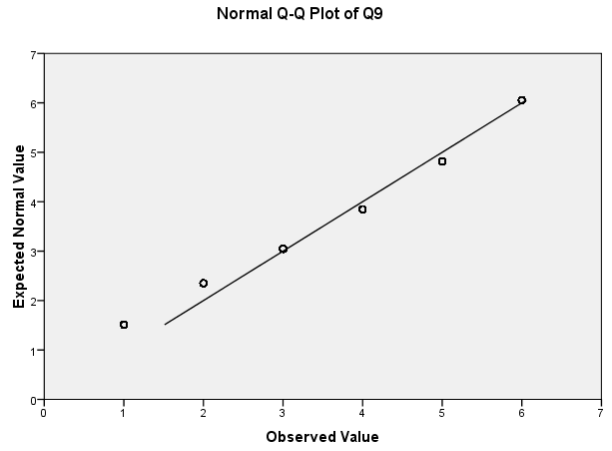
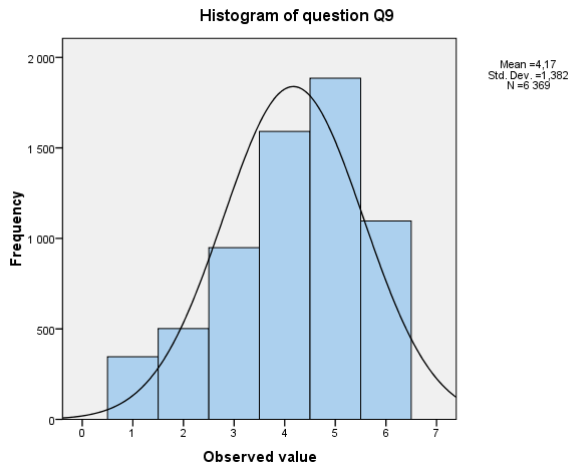
APPENDIX A

In this appendix we present the remaining histograms and normal Q-Q plot of the answers to the survey questions from the distribution analysis.









APPENDIX B

In this appendix we present the remaining histograms and normal Q-Q plot for the key ratios from the distribution analysis.

